Perceptual organization in the tilt illusion

Odelia Schwartz
Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY, USA, The Salk Institute, La Jolla, CA, USA, & Howard Hughes Medical Institute, USA

Terrence J. Sejnowski
The Salk Institute, La Jolla, CA, USA, Howard Hughes Medical Institute, USA, & Division of Biological Science, University of California at San Diego, La Jolla, CA, USA

Peter Dayan
Gatsby Computational Neuroscience Unit, UCL, London, UK

The tilt illusion is a paradigmatic example of contextual influences on perception. We analyze it in terms of a neural population model for the perceptual organization of visual orientation. In turn, this is based on a well-found treatment of natural scene statistics, known as the Gaussian Scale Mixture model. This model is closely related to divisive gain control in neural processing and has been extensively applied in the image processing and statistical learning communities; however, its implications for contextual effects in biological vision have not been studied. In our model, oriented neural units associated with surround tilt stimuli participate in divisively normalizing the activities of the units representing a center stimulus, thereby changing their tuning curves. We show that through standard population decoding, these changes lead to the forms of repulsion and attraction observed in the tilt illusion. The issues in our model readily generalize to other visual attributes and contextual phenomena, and should lead to more rigorous treatments of contextual effects based on natural scene statistics.

Keywords: computational modeling, perceptual organization, structure of natural images

Citation: Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. *Journal of Vision, 9*(4):19, 1–20, http://journalofvision.org/9/4/19/, doi:10.1167/9.4.19.

Introduction

At its most basic, perceptual organization concerns *which* sensory stimuli are treated collectively, and *what* consequences there are for each stimulus in being either aggregated with, or separated from, the others (e.g., Behrmann, Kimchi, & Olson, 2003; Wertheimer, 1923). One central assumption is that the logic of this organization depends on the statistics of natural images, through phylogenetic and/or ontogenetic programming (e.g., Attneave, 1954; Barlow, 1961; Brunswik & Kamiya, 1953; Elder & Goldberg, 2002; Geisler, Perry, Super, & Gallogly, 2001; Krüger, 1998; Sigman, Cecchi, Gilbert, & Magnasco, 2001). In this paper, we consider how the tilt illusion may arise as a downstream consequence of perceptual organization.

In the standard (direct) form of the tilt illusion, the perceived orientation of a central bar of light is tilted *away* from the orientation of a surround grating that forms its visual context (Figure 1a, left). Under particular circumstances, one can also observe a weaker but consistently attractive effect (called the indirect tilt illusion) toward the surround orientation (Figure 1a, right). The tilt illusion is an appealing target for elucidating perceptual organization since its psychological (e.g., Goddard, Clifford, & Solomon, 2008; Wenderoth & Johnstone, 1988; Westheimer, 1990, and see Figures 1a and 1b) and neural (e.g., Cavanaugh, Bair, & Movshon, 2002; Felsen, Touryan, & Dan, 2005; Gilbert & Wiesel, 1990; Li, Thier, & Wehrhahn, 2000; Sengpiel, Sen, & Blakemore, 1997, and see Figure 1c) bases have been extensively probed. The illusion has also been a focus of theoretical interest (e.g., Clifford, Wenderoth, & Spehar, 2000; Gibson & Radner, 1937; Schwartz, Hsu, & Dayan, 2007; Series, Lorenceau, & Frégnac, 2003; Solomon & Morgan, 2006). Indeed, it has long been known that simple, mechanistic, changes to tuning curves at the neural level might underlie observed contextual biases in the tilt illusion (and also in aftereffects, a related phenomenon associated with adaptation, e.g., Jin, Dragoi, Sur, & Seung, 2005; Kohn, 2007; Kohn & Movshon, 2004; Teich & Qian, 2003). However, why the surround stimulus might lead to these tuning curve changes, and the

doi: 10.1167/9.4.19 Received October 4, 2007; published April 24, 2009 ISSN 1534-7362 © ARVO

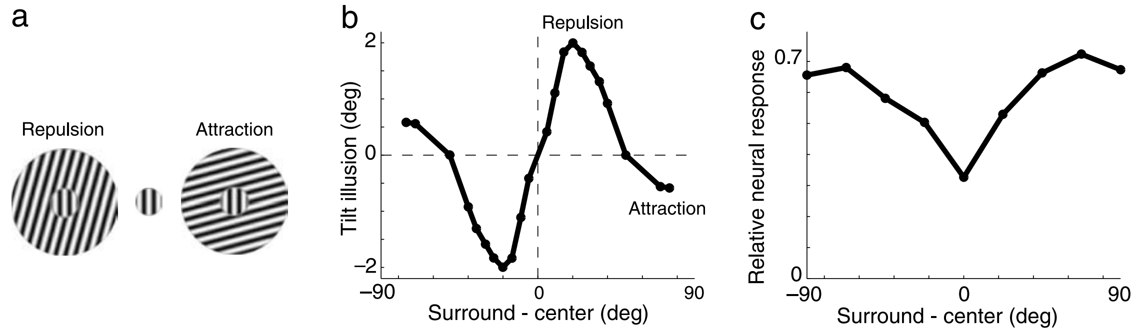

Figure 1. Contextual surround effects. (a) Repulsion and attraction in the tilt illusion. The center grating is vertical; the surround grating makes the center appear repulsed away from the surround (left: surround tilted +15° from vertical); or attracted towards the surround (right: surround tilted +75° away from vertical). For comparison, the grating in the middle is oriented vertically with no surround. (b) Tilt illusion perceptual data from Clifford et al. (2000). The data are an average of 4 subjects in Westheimer (1990). Original data are for positive angles on the *X* axis, and we reflect the data symmetrically for the negative angles, for ease of comparison to simulations. (c) Example electrophysiological data from V1 showing response suppression for neurons associated with the center stimulus as a function of the discrepancy between center and surround orientations. Data replotted from Sengpiel et al. (1997), for 9/32 neurons which showed "iso-oriented suppression."

relationship between this and the statistics of natural scenes, are not clear.

Recent treatments of natural scene statistics focus on the coordination between the outputs of oriented filters at various orientations, scales, phases, and spatial positions (Karklin & Lewicki, 2005; Schwartz, Sejnowski, & Dayan, 2006b; Schwartz & Simoncelli, 2001b; Simoncelli, 1997; Wainwright, Simoncelli, & Willsky, 2001; Zetzsche & Röhrbein, 2001; Zetzsche, Wegmann, & Barth, 1993). In particular, it has been shown that filters with similar orientation preferences representing nearby spatial locations have statistically correlated magnitudes. However, only one main computational implication of this fact has been drawn, namely the redundancy reduction conclusion (Barlow, 1961) that the correlations can be reduced (Schwartz & Simoncelli, 2001b) by a nonlinear model of primary visual cortical neurons, known as divisive gain control (Carandini, Heeger, & Movshon, 1997; Geisler & Albrecht, 1992; Heeger, 1992; Wilson & Humanski, 1993), by dividing out the common magnitude of the correlation. Further, the implications of the coordinated statistics have only been considered for contextual effects at the single neural unit level; not at the level of populations of neurons, nor perception. Moreover, modern versions of the statistical model known as the Gaussian Scale Mixture model (GSM, described in Background to modeling approach section; Andrews & Mallows, 1974; Portilla, Strela, Wainwright, & Simoncelli, 2003; Schwartz et al., 2006b; Wainwright & Simoncelli, 2000; Wainwright et al., 2001), have only been applied in domains outside neuroscience, such as image processing and statistical learning. In this paper we study the implications of the GSM model for orientation surround effects both at the neural population and perceptual levels, with an emphasis on the tilt illusion.

A key question for the model is which filters are statistically coordinated and thereby subject to common divisive gain control. We argue that the direct tilt illusion arises when filter activations in the center and surround locations are coordinated, intuitively because natural visual objects and textures are spatially extended. We demonstrate that the resulting contextual gain control leads to tilt repulsion, through population decoding. However, along with continuous regions that are part of the same object or texture, images also contain regions of segmentation (Blake & Zisserman, 1987; Geisler, 2008; Geman & Geman, 1984; Mumford & Shah, 1989; Shi & Malik, 2000). We suggest that when segmentation is a more likely explanation of the center and surround stimuli, then the gain control pool of the center filter should place less weight on the surround filter activations. We show that this leads to the indirect tilt illusion.

There has been a number of previous theoretical treatments of the tilt illusion or aftereffect (Bednar & Miikkulainen, 2000; Clifford et al., 2000; Dayan, Sahani, & Deback, 2003; Schwartz, Sejnowski, & Dayan, 2006a; Stocker & Simoncelli, 2006; Wainwright, 1999). However, previous approaches have not incorporated findings about the nature of statistical coordination in natural scenes, and in most cases have not encompassed both the neural population and perceptual levels. Our main goal in the paper is not to fit the illusion data better than other modeling approaches, but rather to provide a framework with rich links to modern treatments of natural scene statistics. The key innovation in our approach is combining a recent formal model of natural scene statistics widely used in the image processing and statistical learning literature, with a pre-existing model of inference about (i.e., decoding of) orientation from the activity of a population of neurons. We emphasize that throughout the paper, we are not *learning* the parameters of the GSM model (e.g., as in Schwartz et al., 2006b). The new goal here is rather to study the implications of the GSM framework, for both the neural and perceptual levels. This work sets the basis for future studies incorporating learning (and adaptation) of the model parameters from natural scenes. We expect this framework to be applicable to studying a wealth of other neural and perceptual contextual data (e.g., Clifford & Rhodes, 2005; Li, 1999; Schwartz et al., 2007) in a more rigorous manner based on scene statistics.

We first discuss the background for our approach (Background to modeling approach section). In the Methods section we describe the formal methods and parameters for the neural unit and population decoding model. We next apply the methods to changes in tuning curves and repulsion and attraction in the tilt illusion (Model simulations: Tuning curve changes and the direct tilt illusion section; and Segmentation model and the indirect tilt illusion section). Finally, in the Discussion, we discuss the wider implications of our model and its relationship to the other accounts of the tilt illusion.

Background to modeling approach

As noted in the Introduction, our approach is based on two prominent directions in the literature. The first is a recent model of natural scene statistics, which is closely related to neural gain control. The second is a standard method for decoding neural unit population activity. In this section, we discuss these directions and how they relate to our modeling approach. This is followed by a formal Methods section.

Natural scene statistics and nonlinear neural gain control

Our neural model is a statistically-justified elaboration of a standard divisive gain control model. This section starts with a review of the standard model (Albrecht, Farrar, & Hamilton, 1984; Heeger, 1992), and then proceeds to its elaboration based on a Gaussian Scale Mixture model of natural scenes. The scene statistics and equations are described with the help of cartoons and figures. Note that for completeness, we also include the technical details applied for generating the figures in the captions. However, the actual methods and parameters for the neural unit model are described in detail in the Methods section.

The standard model is formulated at the level of a single neuron or unit in primary visual cortex. We consider a stimulus input consisting of a center and surround stimulus, as in the tilt illusion, and assume that the unit has a receptive field that is spatially overlapping with the center stimulus, and has a given preferred orientation. In physiological terms, the center stimulus is within, and the surround stimulus is outside, the unit's classical receptive field.

According to the standard model, the (nonlinear) response of a unit at the center location is given by dividing its feedforward activation f_c by a gain control signal γ_{gain} (Albrecht et al., 1984; Heeger, 1992):

$$r_c = \frac{f_c}{\gamma_{\text{gain}}}. \tag{1}$$

The feedforward drive $f_c = F_c \cdot I$ is based on a dot product of the unit's tuning function, F_c, sometimes called its filter (c standing for center), with the stimulus, I. The gain control signal γ_{gain} is determined by the rectified (or squared) feedforward activations of other units in center and surround locations. In most models, γ_{gain} also incorporates an additive constant to prevent the denominator from tending to zero. The output of the model is nonlinear due to both the rectification and division. This form of model is often denoted "divisive normalization" or "divisive gain control."

Recent treatments have suggested that this form of divisive model can be motivated from the principle of reducing the redundancy (Barlow, 1961) in the feedforward drives f_c of multiple units that arises because of the particular statistical structure of natural scenes. The main idea is that filters with similar preferences for orientation representing nearby spatial locations in a scene, have striking statistical dependencies; and that divisive gain control can reduce these statistical dependencies (e.g., Schwartz & Simoncelli, 2001b). Figure 2a shows an example of a natural scene taken from the Berkeley database (Martin, Fowlkes, Tal, & Malik, 2001, which we chose for reasons that will become clear later); and Figure 2b shows the characteristic form of the statistical dependency.

Figure 2b plots the joint conditional statistics for the activation of a vertically oriented filter in a center location f_c, given the activation of a non-overlapping vertically oriented filter f_s in a surrounding, contextual, location. The filter activations, f_c and f_s, are computed by convolving each receptive field tuning function (F_c and F_s) with a set of natural scenes (where the receptive fields are spatially displaced relative to one another), and accumulating joint statistics. The figure shows that the magnitudes of f_c and f_s are coordinated in a rather straightforward manner (Schwartz & Simoncelli, 2001b).

We turn this statistical correlation into a gain pool through the medium of what is known as a top-down or generative account (Grenander & Srivastava, 2002; Hinton & Ghahramani, 1997; Neisser, 1967; Zhu & Mumford, 1997). The idea of a generative model is to specify what sorts of statistical structures lead to the filter

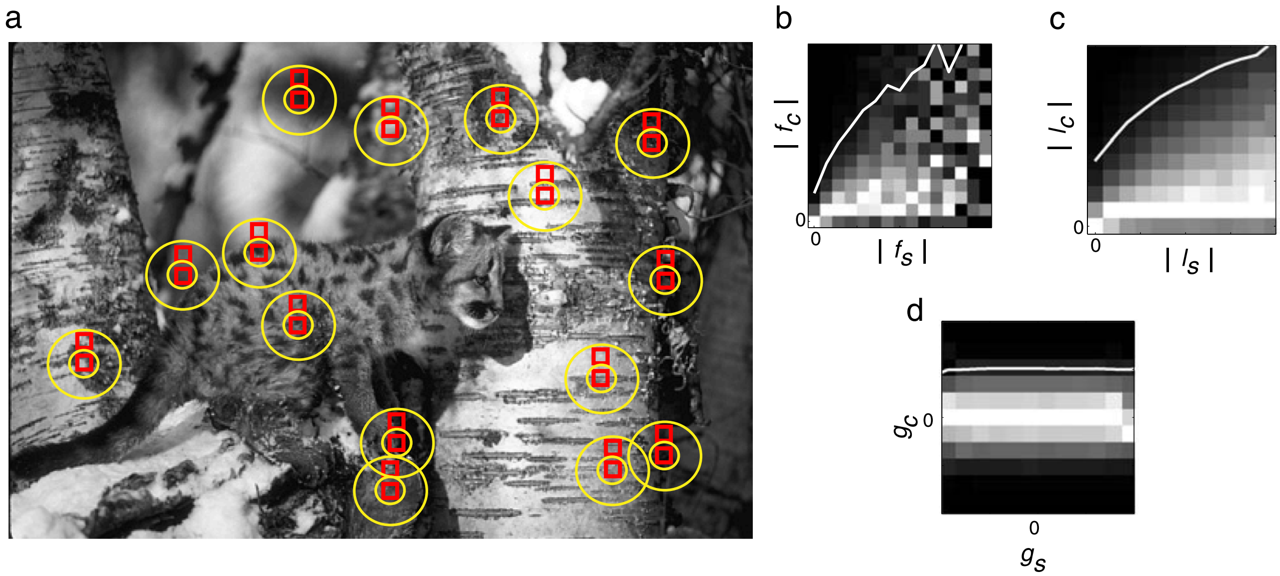

Figure 2. Statistical dependencies in natural scenes. (a) Example image from the Berkeley database (Martin et al., 2001). The concentric yellow circles show example stimulus configurations of center and surround. The center stimulus is the 9 × 9 red square patch in the inner circle; and the surround stimuli are equivalent patches in the outer circle (we mark in red a surround patch just above the center patch). (b) To show the characteristic structure of the joint statistics between center and surround, we gather center f_c, and surround f_s, vertical filter activations. The statistics are collected over 30000 random samples from each of 20 images. Filters are taken from the first spatial frequency level of a steerable pyramid (Simoncelli, Freeman, Adelson, & Heeger, 1992). We plot the joint conditional statistics of the magnitude of activation of a vertical filter $|f_c|$ in the center location, given the magnitude of a vertical filter activation in the surround location $|f_s|$. Intensity is proportional to the bin counts, except that each column is independently re-scaled to fill the range of intensities, giving rise to a conditional distribution. The solid white line marks twice the conditional standard deviation, emphasizing the form of the dependency. (c) Joint conditional statistics of the magnitude of synthetic activations, $|l_c|$ and $|l_s|$, generated according to the GSM model. We generate l_c and l_s, by multiplying a Gaussian variable (g_c and g_s respectively) by a common mixer v. (d) Joint conditional statistics of the Gaussian synthetic variables, g_c and g_s.

activations observed in images. The task for the visual cortex is to invert this model by taking a particular input scene and reporting in the activities of its units what structures might have generated the scene. The structures are *hidden* in that they are not transparently evident in the observed activations of the filters; rather, a (nonlinear) operation must be applied to the image to estimate or recognize them. In our case, we consider a generative model with a particular functional form that gives rise to the observed statistical correlations, creating coordinated values equivalent to f_c and f_s (of Figure 2b). As we will see, the nonlinear recognition operation involves the equivalent of a form of divisive gain control. We propose that the resulting nonlinear model should form the basis for the neural unit model.

To be more concrete, we apply a recent synthetic model called a Gaussian Scale Mixture model (GSM; Andrews & Mallows, 1974; Portilla et al., 2003; Schwartz et al., 2006b; Wainwright & Simoncelli, 2000; Wainwright et al., 2001). The GSM model has been used extensively in image processing and in statistical learning, and its relation to divisive gain control, which we describe below, is also well known. The advantage for us of using this model is that the explicit statistical description of its components provides a firm foundation for exploring the implications of the changing membership of the gain control pool (e.g., the center and surround filters that comprise it) in our neural and perceptual models.

We first describe how to capture statistical coordination in the simplest instance of a GSM model, and then show how this leads to gain control. The model explicitly generates synthetic activations l_c and l_s, that have the same form of statistical coordination as the filter activations f_c and f_s (which are observed in natural scenes between filters corresponding to a center and surround location, as in Figure 2b). According to the GSM:

$$l_c = vg_c. \tag{2}$$

That is, l_c is given by multiplying together two independent random variables. One, g_c, is drawn from a Gaussian distribution, and represents the local form at a point in the image. The other, v, is positive, and is called a mixer. Similarly, the model generates l_s according to:

$$l_s = vg_s. \tag{3}$$

Here g_s is the local form of the surround (another Gaussian variable, which we assume to be independent of g_c), multiplied by the *same* common mixer v. Since the center and surround share a common mixer, the generated filter activations are statistically coordinated. For instance, the samples of g_s and g_c are independent, but if the corresponding sample of v is low, the absolute values of the multiplications $|l_s| = |g_s v|$ and $|l_c| = |g_c v|$ will tend to be low together, and similarly if v is high, these will have a larger probability of being high together. For many such samples, we plot the coordinated statistics of the magnitudes of l_c and l_s generated according to the GSM model (Figure 2c). Note the similarity to the observed empirical statistics of the absolute filter activations $|f_c|$ and $|f_s|$ in Figure 2b.

The connection with divisive gain control is clear, since $g_c = \frac{l_c}{v}$; and $g_s = \frac{l_s}{v}$. Indeed, if we divide our generated l_c and l_s by the common mixer variable, we are left with two Gaussian variables that are independent (Figure 2d). We assume that the goal of a model unit in the center location is to perform exactly this operation, calculating g_c by estimating and dividing out v. This operation depends on the assumption that l_c and l_s are indeed statistically coordinated, an assumption that we will later consider in greater detail.

In practice, as a technical detail, we do not need to first estimate the common mixer and then divide it out; but rather can compute the estimate of g_c directly and analytically, given the linear filter activations l_c and l_s. The expected value $E[g_c|l_c, l_s]$ is taken to be the nonlinear response corresponding to a center model unit, and amounts to a form of divisive gain control. It is given by (Schwartz et al., 2006b, and see also Appendix A for detailed derivation):

$$E[g_c|l_c, l_s] \propto \mathrm{sign}\{l_c\}\sqrt{|l_c|}\sqrt{\frac{|l_c|}{l}}. \quad (4)$$

Here $l = \sqrt{l_c^2 + l_s^2 + k}$ which sets the gain. The term k is a small additive constant, which sets a minimal gain when the filter activations are zero, preventing the denominator from vanishing. We have thus far assumed a single surround filter; if there are more surround filters, then we add their squared activations to the gain. Appendix A gives the full form of Equation 4 and its derivation, including the constant of proportionality, which depends on l and on n, which is the number of center and surround filters comprising l, but not directly on the individual values of l_c and l_s. As we will see in the Methods section, the expectation in Equation 4 will form the basis for our neural unit model.

To illustrate Equation 4 better, Figure 3 shows the GSM model's nonlinear response estimate $E[g_c|l_c, l_s]$ as a function of the surround activation l_s. Specifically, we fix the parameters and the value of l_c in Equation 4, and show how $E[g_c|l_c, l_s]$ changes as a function of l_s. Figure 3 show that the estimated model response $E[g_c|l_c, l_s]$ decreases as l_s increases. This results in surround suppression, whereby a stronger surround strength (and hence a stronger divisive gain control) decreases the response of the unit.

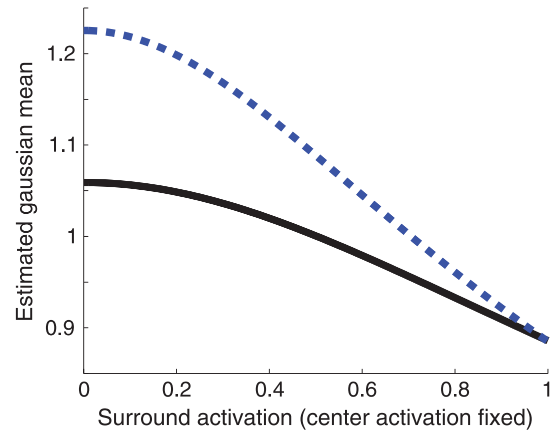

Figure 3. GSM model behavior as a function of surround activation. We plot the estimate of the neural response $E[g_c|l_c, l_s]$ (Equation 4), for a fixed activation of the center (set to $l_c = 1$), varying the surround activation between $l_s = 0$ and $l_s = 1$ (X axis). In the black solid line, the number of surround filters is set to 1 and the number of center filters to 1 (thus overall, $n = 2$); and the additive constant k is set to 0.125. In the blue dashed curve, the number of surround filters is set to 2 and the number of center filters to 1 (thus overall, $n = 3$); and the additive constant k is set to 0.2. A larger surround activation leads to a reduced estimate of the neural response associated with the center.

Of course, l_s only affects g_c if center and surround share the same gain pool. In the model, we set the gain pool to provide the nature of coordination that is actually seen in natural scenes. Analysis of natural scene statistics suggests that filters of similar orientations at different spatial positions will be coordinated to a degree that depends on spatial and orientation differences (Schwartz & Simoncelli, 2001b), an expectation embodied in recent work on the GSM and related models (Karklin & Lewicki, 2005; Schwartz et al., 2006b). Here, we consider this scheme of perceptual organization, and its implications for neural processing and perception in the tilt illusion. We first assume that center and surround filters of the same orientation share a common mixer, and so are in the same gain control pool. Later we relax this assumption.

Unit population decoding approach

We have thus far described a statistical elaboration of a standard nonlinear divisive gain control model, representing a unit at the level of primary visual cortex. We next consider a population of such neural units. We provide a brief overview of the neural population model, for readers who are unfamiliar with the topic, demonstrating the concept of population decoding with the help of a cartoon figure. The detailed methods and implementation of the model are described in the Methods section.

As for a wealth of existing treatments (e.g., Blakemore, Carpenter, & Georgeson, 1970), we assume that the neural population collectively codes for the orientation of a stimulus presented in the center location. First, consider the case of a center stimulus with no surround (Figure 4a, left). Each model unit in the population has a preferred orientation, as depicted in the bottom cartoon of Figure 4a. We can plot the linear filter activations of each unit in the population (Figure 4a, right; in blue). This is peaked and symmetric about the center stimulus orientation. The perceived orientation of the center stimulus can be read out through standard population decoding of the model unit activations (e.g., Dayan & Abbott, 2001; Pouget, Dayan, & Zemel, 2000; Seung & Sompolinsky, 1993; Snippe, 1996). For instance, in the absence of any nonlinear surround influence or noise, a wide range of decoding schemes based on the population activity, including the one we adopt here, the so-called population vector (Georgopoulos, Schwartz, & Kettner, 1986; Jin et al., 2005), will result in a veridical read out.

Now consider an input consisting of both a center stimulus and an annular surround, as in the tilt illusion.

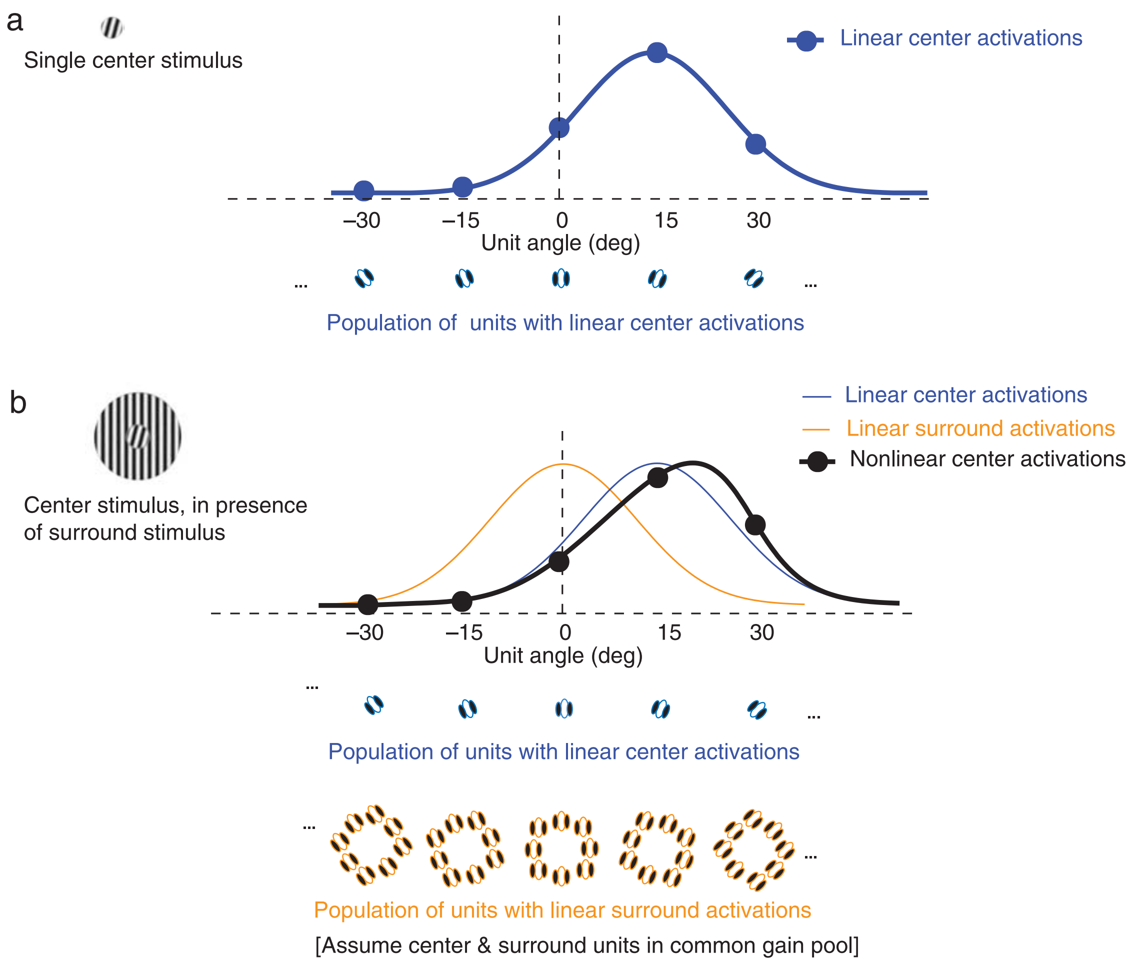

Figure 4. Cartoon of population model. (a) Standard population model in the absence of surround influence. On the left is an example fixed oriented stimulus at 15°, corresponding to a center location with no surround. On the right is a cartoon of a population of primary visual cortical units (each represented by a linear filter with different orientation preference; five example such orientations shown here). The activity of the neural population in response to the input stimulus is just the linear activations of the filters, and these are symmetric and peaked around the center stimulus orientation (in blue; solid circles correspond to the 5 unit orientations on *X* axis). (b) Population model in the presence of surround influence at 0°. On the left is an example fixed tilt illusion stimulus, with both center and surround stimuli. On the right is a cartoon of model units and their activations. We are interested in the *nonlinear* activations of the center units, due to the gain control. We assume each center unit is in a gain pool with surround units preferring its own orientation (see cartoon of 5 center unit linear filters outlined in blue below the *X* axis; and 5 sets of surround unit linear filters outlined in orange below the center units). The linear center activations (blue, thin line) are just the same as in (a); the linear surround activations (orange, thin line) are centered at the angle of the surround stimulus, 0°. The nonlinear center population activity is skewed, and no longer symmetric around the center stimulus orientation, due to the gain control process of the surround activations (black thick curve; solid circles correspond to the 5 unit orientations on *X* axis).

The surround stimulus influences the activities of the units in the population representing the center stimulus. According to our approach, it does so in a nonlinear manner through the gain control process described in Equation 4, due to the statistical coordination between center and surround activations in natural scenes. Specifically, a center unit with a vertical preferred orientation includes in its gain control pool surround units with vertical preferences; a center unit with a horizontal preferred orientation includes activations in the surround with a horizontal preference; and so on (depicted in the bottom cartoons in Figure 4b). The surround units are activated by the surround stimulus according to their orientation preferences (the orange curve), and then, via the gain pools, affect the linear activations of the units representing the center stimulus (the blue curve). This gives rise to a skewed population activity (the black curve) which, via population decoding, leads to a biased perceived orientation involving repulsion (Model simulations: Tuning curve changes and the direct tilt illusion section); and the full tilt illusion (Segmentation model and the indirect tilt illusion section). Figure 4 is intended to provide the abstract intuition and background for our population decoding approach. We now turn to the detailed methods of the model.

Methods

In this section we describe the detailed implementation of the neural units in the population, and the decoding scheme for linking the unit responses to perceptual data in the tilt illusion. The model simulations follow (Model simulations: Tuning curve changes and the direct tilt illusion section; and Segmentation model and the indirect tilt illusion section). For clarity, we describe the methods with regard to the simpler version of model, in which center and surround filters are in the same gain control pool. We later revisit the issue of gain control pools, with a slightly more involved model (Segmentation model and the indirect tilt illusion section).

Neural units

The population model is defined by linear filter activations for each unit corresponding to a center location, together with the nonlinear divisive gain controls for each of those units, determined by both center and surround filter activations. The input in all our simulations is patterned on the tilt illusion, including a center stimulus at one orientation, and a surround stimulus at another orientation.

In detail, we include 360 model units at the central location, with preferred orientations from −90° to +90°. As in most population decoding models, the linear front end of the units (i.e., their input filters) is represented by idealized, identically-shaped, Gaussian tuning curves:

$$l_{ci} = \exp(-(\phi_{ci} - \theta_c)^2 / 2\omega^2), \tag{5}$$

where ϕ_{ci} is the preferred orientation of center unit i in the population; θ_c is a given center stimulus orientation, and $\phi_{ci} - \theta_c$ is the circular difference (in modulo 180, since orientation is circular in 180 instead of 360 degrees) between the preferred orientation of the center unit and the center stimulus orientation. The term ω (set to 22° in our simulations) determines the width of the tuning curves.

As noted earlier, the divisive gain control pool of a center unit is set by the filter in the center location whose linear front-end activation to the center stimulus is denoted as l_{ci} (as in Equation 5), as well as filters in the surround locations, whose linear front-end activations l_{si} in response to the surround stimulus are determined as in Equation 5 according to the difference between the surround stimulus orientation and the surround filter preferred orientation:

$$l_{si} = \exp(-(\phi_{si} - \theta_s)^2 / 2\omega^2), \tag{6}$$

where ϕ_{si} is the preferred orientation of surround unit i in the population; θ_s is a given surround stimulus orientation, and $\phi_{si} - \theta_s$ is the circular difference between the preferred orientation of the surround unit and the surround stimulus orientation. The width of the surround tuning ω is set to 22°, as for the center tuning function. We simplify the treatment of the surround, notably by ignoring its spatial extent and parameterizing surround units only by their preferred orientations. We assume that there are multiple surround filters with preferred orientation ϕ_{si}, each responding equally to the surround stimulus. Inclusion of more surround filters in the simulations will lead to stronger surround influence on the nonlinear gain control. We also assume, based on the statistical coordination described above, that the gain control pool for unit i is set by center and surround filter activations with the same orientation preference $\phi_{si} = \phi_{ci}$.

The nonlinear output of the neural model for each of the center units i in the population is determined by the divisive gain control process. Specifically, the output for each center unit i is given by the GSM model of Equation 4, resulting in estimates $E[g_{ci} | l_{ci};\ l_{si}]$. For convenience, we re-write this equation for each unit i in the population, this time including the constant of proportionality (see Appendix A):

$$E[g_{ci} | l_{ci}, l_{si}] = \mathrm{sign}\{l_{ci}\} \sqrt{|l_{ci}|} \sqrt{\frac{|l_{ci}|}{l}} \left(\frac{\mathcal{B}\left(\frac{n}{2} - \frac{1}{2}, l\right)}{\mathcal{B}\left(\frac{n}{2} - 1, l\right)} \right), \tag{7}$$

where the linear filter activations l_{ci} and l_{si} are given by Equations 5 and 6; the divisive term is $l = \sqrt{l_{ci}^2 + (n-1)l_{si}^2 + k}$, where for our simulations $n = 2$ (because we assume 1 center filter; and the influence of the surround filters is set to 1 is our simulations, which could be thought of as multiple filters with an overall weighting of 1); the additive constant is set to $k = 0.125$; and the last factor in the product (in brackets) involves what is known as a modified Bessel function of the second kind (see also Grenander & Srivastava, 2002).

Finally, for readability, we will denote the output estimate of unit i in Equation 7 as simply g_i, thus dropping the expected value and the c subscript corresponding to center (since the perceptual task only demands decoding at this center location).

To summarize, each center unit in the population has a linear tuning according to Equation 5 and a nonlinear gain control according to Equation 7. The gain control is determined by filters in center and surround with similar orientation preference. The free parameters we set for the simulations are the linear function tuning widths ω; the number of center and surround filters n; and the additive constant k. These parameters affect the strength of the results, but not their qualitative nature.

Population decoding

In order to turn the final activities g_i of the population of center units into an estimate of the angle of the center stimulus, we use a standard population vector decoding scheme (Georgopoulos et al., 1986). According to this, the center angle is given by

$$r = \frac{1}{2}\text{angle}\left\{\sum_i g_i \mathbf{u}(2\phi_i)\right\}, \tag{8}$$

where ϕ_i is the preferred angle of unit i, and $\mathbf{u}(\phi)$ is a two dimensional unit vector pointing in the direction of ϕ (and the doubling takes account of the orientation circularity).

Model simulations: Tuning curve changes and the direct tilt illusion

We now apply the methods to study the changes in the tuning curves and perception in the tilt illusion due to spatial context. We examine the effect of spatial context by varying the orientation of a central stimulus in the presence and absence of a surround stimulus (which is presented at the fixed orientation of 0°). First, we consider the effect the surround has, via gain control (Equation 7), on the tuning curves of the units coding for the center (Equation 5). Then, we examine how these changes to the tuning curves give rise to the tilt illusion through decoding the population activity (Equation 8).

Changes in tuning curve

Figure 5a shows the activity of a few sample units in the population as a function of center orientation in the absence of a surround stimulus. Two of these tuning curves (red: 0°; green: 45°) are picked out. By design, these have the same shape, since the gain controls (from Equation 7, with $l_{si} = 0$) affecting the linear activation of these units (from Equation 5) are the same.

Figure 5b shows the consequence of presenting a surround stimulus (at 0°), for the tuning curves associated with the center. The surround reduces the peak height of the tuning curve when the preferred orientation of the unit is equal to the surround stimulus orientation (Figure 5b; red), but not when the preferred orientation is far from the surround (Figure 5b; green). This is due to the gain control, which, according to our rules of perceptual organization is computed over surround units with similar orientations. For a unit with preferred orientation equal to the surround, the gain control signal is larger, resulting in more suppression of the response (see also Figure 3). The reduction in tuning curve height when the surround orientation is near its preferred orientation is widely documented in physiological data (e.g., Cavanaugh et al., 2002; Sengpiel et al., 1997, and see Figure 1c).

According to Figure 5b, the tuning curves of model units whose preferred orientations are very close to that of the surround also narrow slightly. Some physiological studies have found changes in tuning widths (Gilbert & Wiesel, 1990), but the trend as a function of preferred and surround orientations is inconclusive. Some experimental data also show repulsive shifts in the preferred orientations of the cells (e.g., Felsen et al., 2005; Gilbert & Wiesel, 1990), although it has been suggested that the shifts might be a consequence of underestimating the size of the classical receptive field experimentally, and hence of the surround actually impinging upon the receptive field (Cavanaugh et al., 2002).

Changes in tilt perception

We derive the perceptual consequences of these effects on the tuning curves by decoding the population activity of the center units. Figure 6a shows the full population activity to a 20° center stimulus in the presence (filled, black) and absence (open, blue) of a 0° surround. Without a surround, the population response is symmetric about 20°, and population vector decoding (Equation 8) leads to unbiased inference of the center angle (open, blue arrow on the X axis). With the surround, the response is asymmetric about 20° because of the suppressive effects

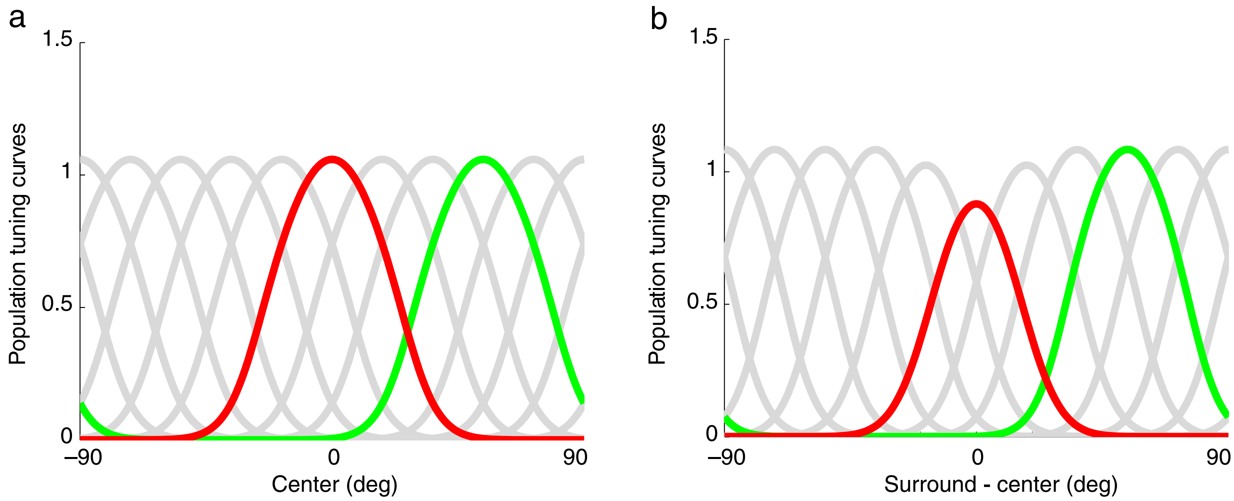

Figure 5. Changes in tuning curves in the GSM model. (a) Tuning curves in the absence of a surround stimulus (l_{si} = 0 for each unit *i*), plotted as a function of the angle of the center stimulus. The center filter activations l_{ci} are computed for each preferred orientation in the population according to Equation 5; the nonlinear neural activation is then given by $E[g_{ci}|l_{si}, l_{ci}]$ according to Equation 7. The red curve corresponds to the center preferred orientation 0°; the green curve to 45°; other example tuning curves are shown in gray. (b) Tuning curves in the presence of a surround stimulus (with a fixed orientation of 0°), plotted as a function of the difference between the angle of the surround and center stimulus. Both l_{ci} and l_{si} are computed according to Equations 5 and 6. The resulting nonlinear activation is then given by Equation 7. Color scheme is the same as in (a).

of the gain control on the units near 0°. Using population vector decoding, the inferred angle is 22.4° (filled, black arrow on the *X* axis). The difference between the perceived and presented angle of the center stimulus is therefore 2.4°, repulsively biased away from the 0° surround. This realizes the direct tilt illusion (e.g., Wenderoth & Johnstone, 1988).

Figure 6b shows the differences between the inferred and actual center angles in the presence of the surround, as a function of the center angle. When center and surround are the same (0°), the perceived orientation is unbiased; when the surround stimulus has a different orientation from the center stimulus, the perceived orientation is repelled from the surround orientation.

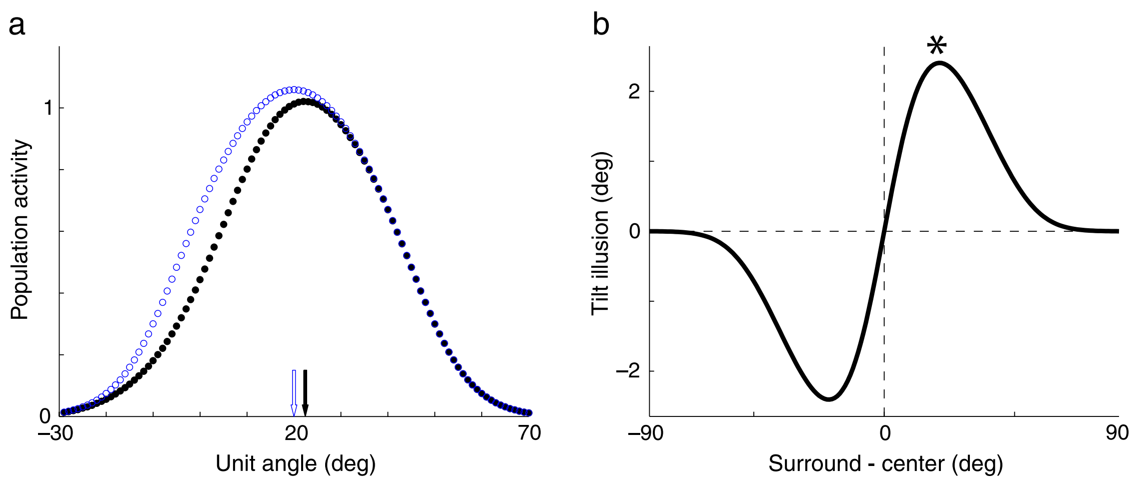

Figure 6. Population activity and decoding for GSM model. (a) Population activity for a fixed 20° center stimulus in the presence (filled, black) and absence (open, blue) of a 0° surround. The *X* axis corresponds to the preferred orientation of each unit in the population, and the *Y* axis to its activation in response to the stimulus. Note that we have re-centered the *X* axis around the unit with a preferred angle equal to the center stimulus angle (20°). The inferred angles are computed using population vector decoding, and are marked with arrows (open blue arrow: no surround; filled black arrow: 0° surround). (b) Bias in the tilt illusion resulting from population vector decoding, as a function of the difference of the angle of the surround and center stimulus. The surround is fixed at 0°. The bias is given by the difference between the perceived and presented angles of the center stimulus. The star marks the condition plotted in (a) for a center stimulus at 20°.

Repulsion is greatest for relatively small (20°) differences between the center and surround angles. The model simulations (Figure 6b) show a similar trend to experimental data (Clifford et al., 2000, and the repulsion in Figure 1b). Note that previous work has pointed out that the kind of changes in tuning curve observed in Figure 5b (and in particular, the reduction in tuning curve heights when the preferred orientation is similar to that of the surround) will lead to repulsion through standard population decoding (e.g., Gilbert & Wiesel, 1990; Jin et al., 2005; Wilson & Humanski, 1993). Here we have shown that our model of natural scene statistics achieves these sorts of changes in tuning curve.

Segmentation model and the indirect tilt illusion

Figure 6b clearly indicates that the model only exhibits repulsion, and thus does not reproduce the indirect tilt illusion, where for large angle differences between center and surround, one observes attraction toward the surround orientation (as in Figure 1b). Arguably, the indirect effect is more difficult to explain, and has been addressed in only a restricted set of the previous functional models of the tilt aftereffect (e.g., Bednar & Miikkulainen, 2000; Clifford et al., 2000). In this section, we argue that the model has over simplified the segmentation that is appropriate to natural scene statistics. Rectifying this, which implies a more sophisticated account of gain control, leads to a unified account of direct and indirect tilt illusions.

We begin by revisiting the issue of statistical coordination of filter activations in center and surround locations in natural scenes. The gain control model of Equation 7 is based on the statistical observation that the magnitudes of spatially displaced filters of similar orientation are correlated for typical natural scenes (e.g., Figure 2b). According to this model, a common gain signal is computed from filters preferring the same orientation that are activated by the center and surround stimuli. However, Figure 7 points out an important fact that we have so far omitted in our model; namely that this coordination only applies when center and surround stimuli are actually part of the same visual segment or object. Indeed, we used the Berkeley database since hand-labelings of segmentation boundaries are available (Martin et al., 2001). Figures 7a and 7b show the statistical relationship between vertically-oriented filters taken from patches of scenes that were classified by human subjects as belonging to the same segment (*within*); the statistics follow the same correlation structure as in Figure 2b. By contrast, Figures 7d and 7e show the statistical relationship between vertically-oriented filters taken from patches of scenes *across* segmentation boundaries. It is apparent that the coordination is greatly reduced, implying that the surround filter does not provide evidence about the mixer (v) for the center, and therefore should not form part of the gain pool for the center. This statistical coordination within versus across segments in natural scenes has not been previously shown empirically, nor incorporated in neural models. As before, we have included in the caption of Figure 7 the technical details for creating the scene statistics figure, but these are not critical for understanding the methods for the simulations which follow below.

These statistical observations lead to three key questions. First, how should the model decide whether center and surround stimuli are part of the same visual object? Second, what is the effect on the tuning curves associated with the center stimulus of this decision? Both the decision and the effect are determined by the relative orientations of center and surround stimuli. Third, what are the perceptual consequences of these effects on the tuning curves? The decision about whether the surround stimulus should be considered as forming part of the same gain pool as a unit representing the center stimulus was the topic of some detailed modeling in Schwartz et al. (2006b). In the full model, this decision depends on the whole collection of activities of all the filters across the whole input stimulus; that is, on l_{ci} and l_{sj}. However, the basic effect, which we use in a simple and abstract form, is the obvious one that the closer the center and surround angles, the more likely they are to be part of the same visual object. Figures 7c and 7f show a coarse version of this distribution for the *within* and *across* conditions taken from the Berkeley segmentation data themselves.

Based on the statistical observations in the Berkeley image segmentation database, we now describe the methods for the revised gain control model. In our simple model of this, a unit i with preferred angle φ_{ci} representing the center stimulus, treats a surround stimulus of angle θ_s as being part of the same visual object with probability:

$$p = \exp(-(\phi_{ci} - \theta_s)^2 / 2\lambda^2), \tag{9}$$

where λ is a parameter that controls the steepness of this selection (set to $\lambda = \sqrt{4000^\circ} = 63.2456°$ in our simulations).

If the surround stimulus is not treated as being part of the same gain pool as the center stimulus, then its activity does not bear on the gain signal for the center, and so the activity of unit i would be $E[g_{ci} | l_{ci}]$, given just the observation at the center l_{ci} (i.e., the gain signal is only comprised of the center filter activation l_{ci}). If center and surround are coordinated, then, as before, the activity would be $E[g_{ci} | l_{ci}, l_{si}]$, given both observations (i.e., the gain signal is given by both the center filter activation l_{ci} and the activation of the surround filters l_{si}). We assume that the net response is just the average of

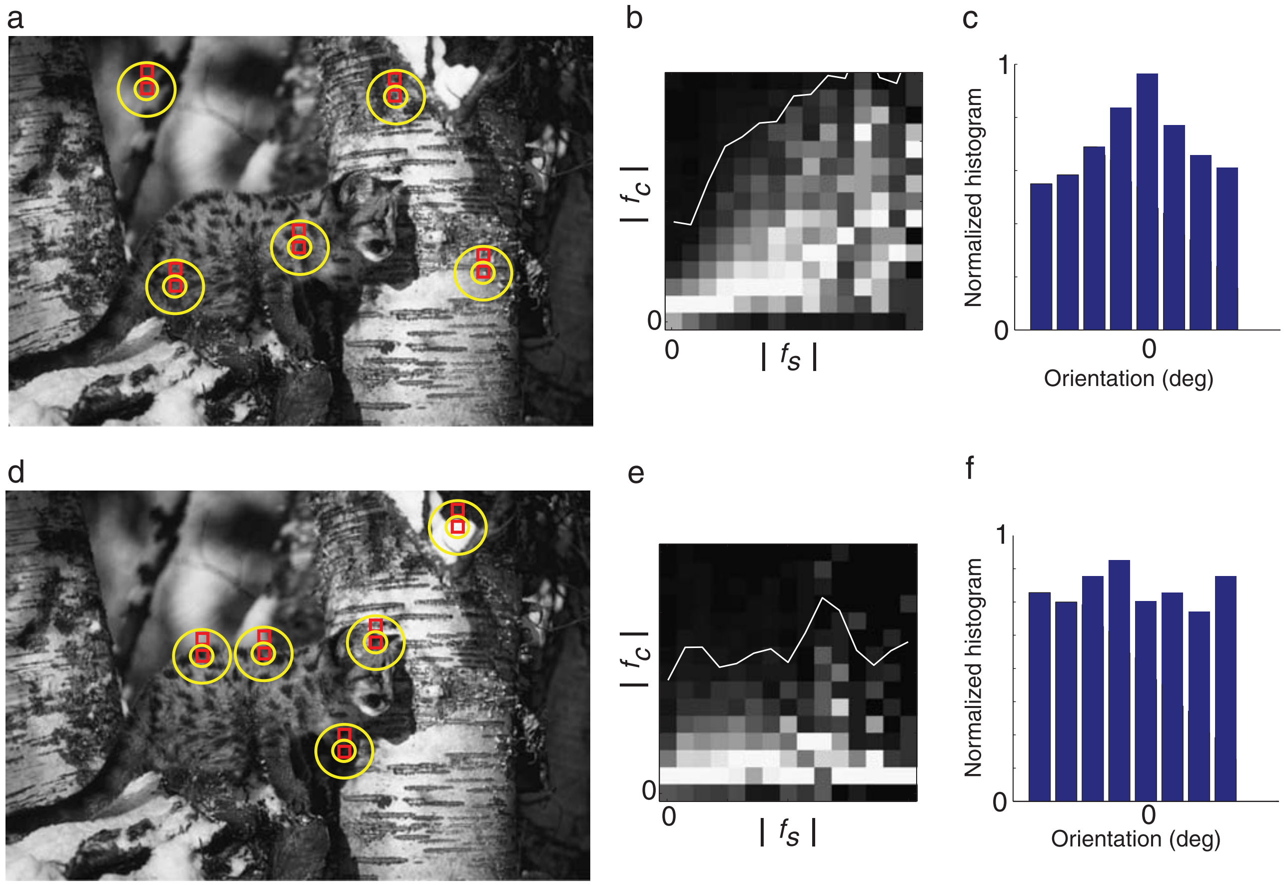

Figure 7. Statistical dependencies in natural scenes *within* (a–c) and *across* (d–f) boundaries. (a–c) Example image from the Berkeley database (Martin et al., 2001) and statistics of patches *within* portions of the scene classified by at least two subjects as belonging to the same segment. (d–f) Example image from the Berkeley database and statistics of patches *across* hand-labeled segmentation boundaries. The numbers of samples are matched in the two conditions. *Left column*: Example image and patches corresponding to the *within* (a) and *across* (d) conditions, based on labelings of the first two subjects in the database. We consider patches to be *across* if at least 71 of the 81 pixels in the patches are labeled as belonging to different segments. The red squares correspond to a center location, and a surround location that is above the center. *Middle column*: Joint conditional statistics of the magnitude of a vertical filter activation $|f_c|$ in the center location, given the magnitude of a vertical filter activation in the surround location $|f_s|$. *Right column*: Histogram of the orientation of the center patch, conditional on the surround patch orientation being vertical (0°). We normalize this histogram by the marginal orientation histogram, to avoid biases due to the over-representation of cardinal axes in the patches. The orientation of a patch is determined by computing the filter activations of 8 possible orientations and choosing the maximum above a minimal threshold.

these two means, weighted by their respective posterior probabilities:

$$pE(g_{ci}|l_{ci}, l_{si}) + (1 - p)E(g_{ci}|l_{ci}). \qquad (10)$$

Otherwise, the methods are exactly as before. In particular, the linear front-end tuning curves are idealized (see Methods section, and Equations 5 and 6), and the nonlinear gain control estimation is given according to Equation 7. The remaining parameters are also set according to the Methods section.

Changes in tuning curve

We apply the revised model to simulating tuning curve changes and the tilt illusion. Figure 8 shows the (non-linear) neural activity of a few sample units in the population as a function of center orientation, either in the absence (Figure 8a) or presence (Figure 8b) of a surround (at 0°). The neural responses are computed using Equation 10, with the weights p set according to Equation 9. As before (compare to Figure 5b), the presence of the surround stimulus reduces the peak height of the tuning curve when the preferred orientation is equal to that of the surround. However, when the preferred orientation is farther from the surround, the response is actually enhanced.

This enhancement arises since:

i. the center units share gain pools with surround units having the same preferred orientations;
ii. when the surround stimulus is far from the center stimulus, it only weakly activates the surround units

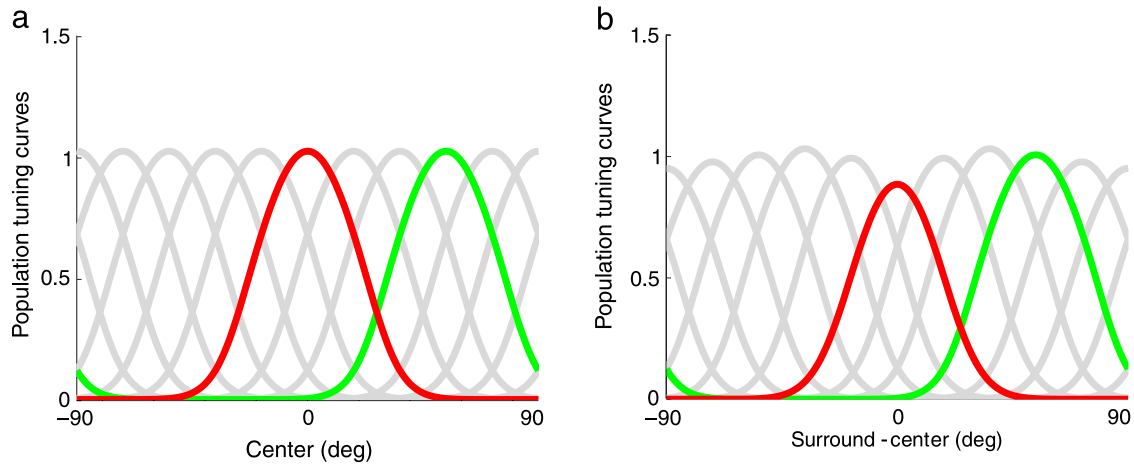

Figure 8. Changes to the tuning curve under the segmentation version of the GSM model. (a) Tuning curves in the absence of a surround stimulus (l_{si} = 0 for every unit i), plotted as a function of the angle of the center stimulus. The center filter activations l_{ci} are computed for each preferred orientation in the population according to Equation 5. Red curve corresponds to preferred orientation 0°; green curve to 45°; and other example tuning curves in gray. (b) Tuning curves in the presence of the surround (whose orientation is fixed at 0°), plotted as a function of the difference between the angle of the surround and center stimulus. Both l_{ci} and l_{si} are computed according to Equations 5 and 6. The color scheme is equivalent to (a).

with similar preferred orientations to the most activated center units;

iii. critically, if a surround unit has a weak activation and is in the gain pool for a center unit, then it provides evidence that the common mixer variable v is small, and that therefore the g_{ci} underlying l_{ci} is large.

If it is not in the gain pool, then its weak activation does not provide any evidence about the mixer v for the center, and therefore does not boost g_{ci}. This boosting therefore waxes as the difference between center and surround orientations grows, but then wanes, since the weighting term, p, gets smaller when the difference is very large. Although this trend has not been documented systematically in neural data, boosting for far angle separations of center and surround has been observed physiologically (e.g., Levitt & Lund, 1997). Interestingly, the mechanistic disinhibition model of (Dragoi & Sur, 2000) shows a similar trend of waxing and waning.

Changes in tilt perception

Finally, we consider the perceptual consequences of this boosting. Figure 9a, which should be compared with Figure 6a, shows the full population activity to a 70° center stimulus in the presence (filled, black) and absence (open, blue) of a 0° surround. Without the surround, the population response is symmetric about 70°, and population vector decoding (Equation 8) leads to unbiased inference of the center angle (open, blue arrow on the X axis). With the surround, the response is asymmetric about 70° (but in exactly the opposite direction of Figure 6a), leading to an inferred angle of 69.41° (filled, black arrow on the X axis). The difference between the perceived and presented center angle is now −.59°. This attraction to the surround orientation is an example of the (weaker) indirect tilt illusion (e.g., Wenderoth & Johnstone, 1988).

Figure 9b shows the differences between the inferred and actual center angles in the presence of the surround as a function of the center angle (compare to Figure 6b). Repulsion is still greatest for small (20°) differences between the center and surround angles. In addition, for large (e.g., 70°) differences between these angles, the model leads to attraction. The repulsion is, as before, a consequence of center and surround activations contributing to the gain pool when the preferred orientation is close to the surround stimulus orientation. By contrast, the attraction is due to the additional force of the segmentation of gain pools.

It is important to note that while the qualitative effects of the tilt illusion (e.g., the direct and indirect effects) are typically significant in different data sets, the quantitative effects (e.g., the strength of the effects) are variable. The strength of reported effects for both the illusion (e.g., Goddard et al., 2008; Over, Broerse, & Crassini, 1972; Wenderoth & Johnstone, 1988; Westheimer, 1990) and aftereffect (e.g., Clifford et al., 2000; Gibson & Radner, 1937; Muir & Over, 1970) have each ranged between approximately 4° to below 1.5° for the direct effect; and a smaller indirect effect typically around 1° or below. For instance, in the illusion data we used which was an average over 4 subjects (Clifford et al., 2000), there is variability across subjects (Westheimer, 1990). The strongest effects for one of the subjects was approximately a direct effect of near 4° and an indirect effect of over 1°, while the weakest effect was closer to 1.25° direct and an almost absent indirect effect. The strength of effects are also influenced by the experimental design, and issues such

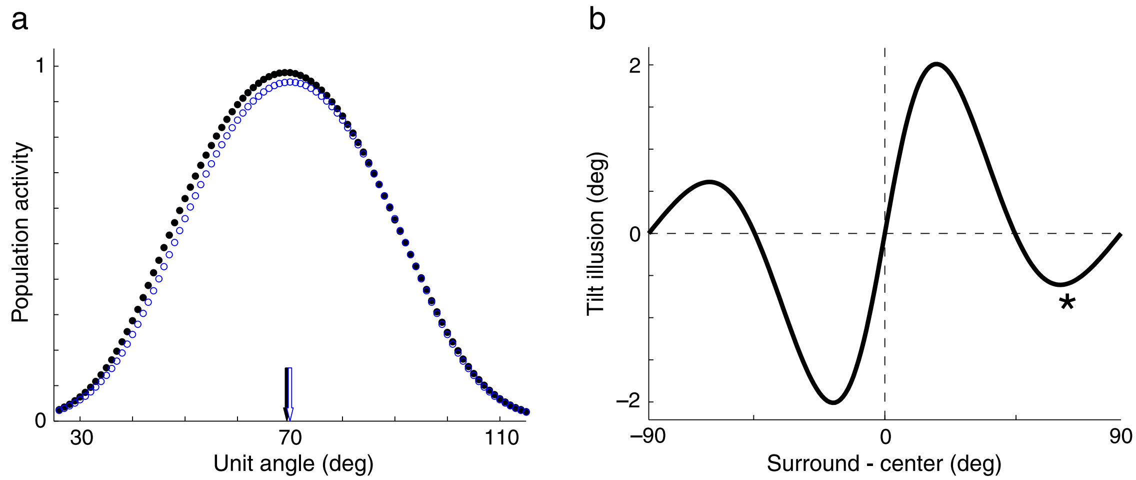

Figure 9. Population activity and decoding under the segmentation version of the GSM model. (a) Population activity for a fixed 70° center stimulus in the presence (filled, black) and absence (open, blue) of a 0° surround. The *X* axis corresponds to the preferred orientation of each unit in the population, and the *Y* axis to its activation in response to the stimulus. Note that we have re-centered the *X* axis around the unit with a preferred angle equal to the center stimulus angle (70°). Also, note that orientation is circular modulo 180, so the label of 110° on the *X* axis is equivalent to −70°. The inferred angle is computed using population vector decoding, and marked with an arrow on the *X* axis (open blue arrow in the absence of surround; filled black arrow in the presence of surround). (b) Bias in the tilt illusion resulting from population vector decoding, as a function of the difference between the angle of the surround and center stimulus (for a fixed 0° surround). The bias is given by the difference between the perceived and presented center orientations. The star marks the condition plotted in (a) for a center stimulus at 70°.

as the stimulus presentation time (Wenderoth & van der Zwan, 1989; Wolfe, 1984), type of stimulus, and so on.

In our model, the strengths of attraction and repulsion are determined by the potency of the effects of the gain control on the tuning curves. Changes in the heights of these curves are more pronounced as the surround strength increases (recall that we chose a surround to center weighting of 1:1 in the simulations). In addition, a steeper probability fall-off of p in Equation 10 leads to a stronger waxing and waning of tuning curve height for surround orientations far from the preferred orientation of the neuron (of Figure 8), and thus to stronger attraction. Also, the orientations at which these happen are influenced by the widths of the idealized Gaussian tuning curves in Equations 5 and 6 (which, as noted in the Methods section, were chosen as 22° in the simulations). Changes in the tuning widths themselves due to the GSM model are more pronounced as the additive constant k in Equation 7 decreases. We do not obtain shifts in tuning curve, although repulsive shifts could arise by incorporating mechanisms similar to the neural adaptation model in (Wainwright, Schwartz, & Simoncelli, 2002). The

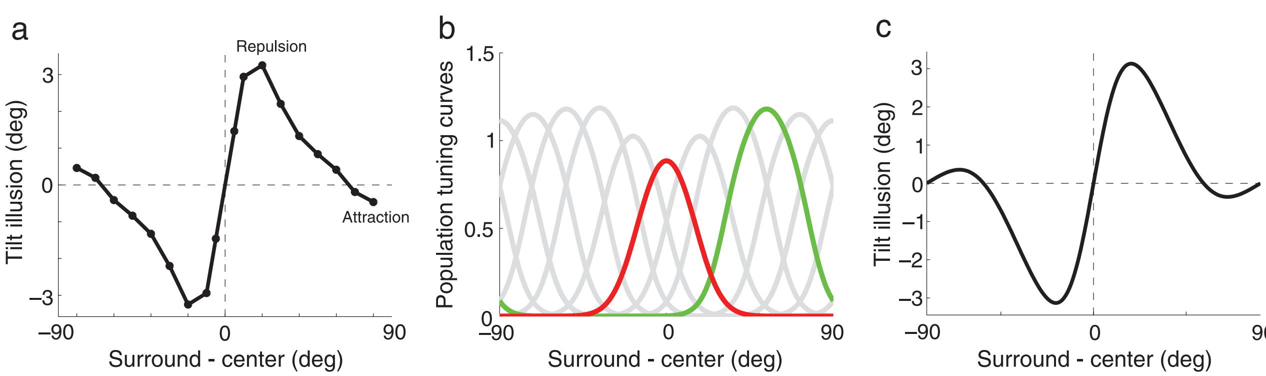

Figure 10. Segmentation model for Goddard et al. illusion data. (a) Tilt illusion data for grating stimuli, for an average of subjects from Goddard et al. (2008), supplementary data). Original data are for positive angles on the *X* axis, and we reflect the data symmetrically for the negative angles, for ease of comparison to simulations. (b) Model tuning curves in the presence of a surround stimulus, plotted as a function of the difference between the angle of the surround and center stimulus. (c) Model bias in the tilt illusion resulting from population vector decoding, as a function of the difference between the angle of the surround and center stimulus. The bias is given by the difference between the perceived and presented center orientations.

strength of the bias in attraction and repulsion is also influenced by the choice of decoding method (Jin et al., 2005). Although the qualitative results are similar, as indeed might be expected, a decoding rule based on the maximally activated unit in the population leads to stronger biases than the population vector method.

Figure 10 shows another example illusion set (Goddard et al., 2008), and our model simulation to accommodate the data. The model parameters are set to $n = 3$ and $k = .2$ (see also Figure 3, dashed line); center and surround tuning widths of 20° and 27° respectively; and $\lambda = \sqrt{10500^{\circ}} = 102.4695^{\circ}$. Note that in addition to the different strength of the direct and indirect effect, the angle of cross-over on the X axis from repulsion to attraction is different. We account for this in the model with the change in tuning widths of the front-end center and surround tuning functions.

Discussion

We described a model for the perceptual organization of visual orientation, applied to the well known tilt illusion. Our main goal was to provide a framework based on modern treatments of natural scene statistics, and to apply this framework across neural and perceptual levels of analysis. In our model, changes in tuning curves at the neural unit level were set according to a recent model of natural scene statistics known as the GSM that is closely related to divisive gain control. The tuning curve changes, in turn, resulted in the tilt illusion through population decoding.

The tilt illusion and related phenomena have been of ongoing interest from experimental and modeling perspectives (for overviews, see, e.g., Clifford et al., 2000; Schwartz et al., 2007, and references therein). Mechanistic population decoding models have shown classes of contextually-mediated changes in tuning curves that can bring about repulsion and attraction (e.g., Gilbert & Wiesel, 1990; Jin et al., 2005; Kohn & Movshon, 2004; Teich & Qian, 2003). However, the mechanistic models do not explain *why* these tuning curve changes occur in the face of contextual stimuli. Functional models of tilt perception have included variants of efficient coding (e.g., Bednar & Miikkulainen, 2000; Clifford et al., 2000; Wainwright, 1999) and Bayesian approaches (e.g., Dayan et al., 2003; Stocker & Simoncelli, 2006), some, but not all, of which can accommodate attraction as well as repulsion. Repulsion and attraction have also been modeled in other domains, such as disparity (e.g., Lehky & Sejnowski, 1990). The main differences between our framework and previous functional modeling approaches are that (a) our model is based on a modern treatment of statistical coordination in natural scene statistics, whereas few previous frameworks have been tied to natural scene statistics models at all; and (b) our model encompasses both the neural unit level tuning curve changes and the effects of these on perception.

Clifford et al. (2000) suggested a model in which the whole population of neurons is shifted and scaled such that the collection of responses remains circularly symmetric about the origin, reducing the redundancy among the neurons. They explained how these changes at the population level can come about through changes in tuning curves.

Clifford's framework offers an appealing account of both the direct and indirect tilt illusion; however, it is not formulated to make connection to the statistics of natural scenes. Another interesting model is that of Bednar and colleagues (Bednar & Miikkulainen, 2000; Sirosh & Miikkulainen, 1997), who suggested a mechanistic self-organizing network, with inhibitory long range lateral connections and short-range lateral excitation. Their model sparsifies and decorrelates the input, leading to both a direct and indirect effect. The indirect effect in their model arises from the network consideration that the total inhibitory connection strength to each neuron is limited. Their model is related to natural scenes ideas, but not at the level of the prominent statistical high-order magnitude correlations on which we focused. Their model is also formulated to explain perception in the tilt aftereffect (not surround contextual effects in the tilt illusion), and does not examine the neural unit tuning curve changes.

We focused on a statistical elaboration of a standard gain control model, which has a rich mechanistic and functional pedigree (e.g., Geisler & Albrecht, 1992; Graham & Sutter, 2000; Heeger, 1992); indeed, Wilson and Humanski (1993) have even shown that a population model of gain control could explain perceptual repulsion. However, by contrast with this work, the functional basis of our framework is based on a statistical generative model of natural scenes, namely the GSM (Andrews & Mallows, 1974; Portilla et al., 2003; Schwartz et al., 2006b; Wainwright & Simoncelli, 2000; Wainwright et al., 2001). Previous gain control models of natural scene statistics have been applied only at the cortical neural level (Schwartz & Simoncelli, 2001b; Wainwright et al., 2002), and the implications for population coding and perception have not been explored.

In addition, previous gain control models have focused on a single gain control pool. We explored a more sophisticated statistical model of gain control, which explicitly allows for both grouping and segmentation of the center and surround gain pools. We motivated the segmentation model using the labeled Berkeley segmentation database, by showing empirically that the coordination of filter activations in center and surround locations depend on whether these locations form part of the same segment within the scene (Figure 7). We showed how this proposal can lead to suppression and boosting of the neural response; and to repulsion and attraction in the tilt illusion. In the first instance, we assumed that center and surround filters of the same orientation always share a common mixer, and so are in the same gain control pool. This led to the direct tilt illusion (Model simulations: Tuning curve changes and the direct tilt illusion section). We later relaxed this assumption, and suggested that to the extent that center and surround are likely to correspond to different segments, the gain control pool of the center should place less weight on the surround filter activations. This consideration allowed the model to account for the indirect tilt illusion (Segmentation model and the indirect tilt illusion section). Indeed, one could perhaps interpret this in terms of Wenderoth and Johnstone (1988); Wenderoth and van der Zwan's (1989) suggestion that the direct and indirect tilt illusions depend on different structural features of the stimuli.

Our version of the GSM is flexible in allowing center and surround filters to have a probability less than 1 of belonging to the same gain pool. This notion of separate gain control pools is not new in neurophysiology; for instance, it has been shown that the activities of V1 neurons are well accounted for by fitting a gain parameter determined for the center and a gain parameter determined for the surround (Cavanaugh et al., 2002). We suggested that when center and surround are more likely to belong to different segments of the scene, the gain control pool of the center should depend less on the surround activations. The assignment of center and surround to gain pools and the evaluations of the gains (i.e., the mixer values) are interdependent, since the values of the mixer variables influence the segmentation (Li, 1999; Zhaoping & Jingling, 2008). Inferring both together in general cases is therefore tricky. For the case of orientation filters and the GSM model, we have previously shown in Schwartz et al. (2006b) (see also Williams & Adams, 1999) a probabilistic solution for this inference. Here, given the austere nature of the scenes involved in the tilt illusion, we used an algorithmic simplification.

In our implementation, the decision about grouping and segmentation was pre-determined. However, an important direction for future work is to learn this grouping and segmentation from the statistics of ensembles of natural scenes; and indeed, there are natural accounts of learning statistical coordination to which to turn (Karklin & Lewicki, 2005; Schwartz et al., 2006b). Further, in this paper we assumed that filters with similar orientations in center and surround locations should be statistically coordinated and hence share a common gain control pool. Along with this basic similarity effect, recent work has also reported more diverse groupings that can include multiple orientations at a given location; geometrical aspects of the organization; and grouping according to spatial frequency. We would expect that these richer groupings would lead to a range of more subtle context based illusions, and this could form the basis of an experimental test of our proposal.

In our model, the relative importance of center and surround, and thus the nature and magnitude of the tilt illusion, is strongly influenced by assumptions about noise or signal (i.e., visual contrast). This is another pressing direction for experimental investigation, particularly because there are competing possibilities for the effects (Zhaoping & Jingling, 2008). For instance, if the contrast of the center is low, then the orientation prevalent in the surround can provide important constraints on the orientation at the center. However, large contrast variation is also a cue to segmentation (e.g., Durant & Clifford, 2006), which would weaken the influence of the surround (Zhaoping & Jingling, 2008). There are actually various ways to incorporate noise in GSM models, including the direction explored by (Portilla et al., 2003) in their work on cleaning up degraded images. The relative contribution of center and surround to the neural gain pool might also be influenced by aspects of attention (e.g., Sundberg, Mitchell, & Reynolds, 2006), another natural extension of our framework.

Grouping and coordination, and their effects on sensory inference, are, of course, not confined to tilt. Illusions in the same family as the tilt illusion have been widely documented for spatial frequency, lightness, motion, color, blur, and faces, as well as other modalities (see references in Schwartz et al., 2007), and it would be interesting to extend our framework to them. Patterns of statistical coordination have duly been found in other domains such as spatial frequency (e.g., Karklin & Lewicki, 2005), and sound (Karklin & Lewicki, 2005; Parra, Spence, & Sajda, 2000; Schwartz & Simoncelli, 2001a). The cases of lightness and color are particularly promising, since many accounts of color constancy and brightness illusions are based on inference in the face of the multiplication between a local reflectance structure and a more global illuminant (e.g., Adelson, 2000; Gilchrist et al., 1999; Nundy & Purves, 2002; Zhaoping, 2007).

From a physiological perspective, the gain control framework of grouping and segmentation could be applied to mid-level vision structures. For instance, it has been suggested that motion area MT receives divisively normalized V1 activations as its input (Rust, Mante, Simoncelli, & Movshon, 2006; Simoncelli & Heeger, 1998), which could then undergo an additional stage of linear weighting and gain control. The possibility of gain control groups, and segmentation according to properties such as depth and layers, have not been extensively explored in this context. This framework might also be applicable to the ventral visual pathway, for pooling of additional invariances (e.g., Cadieu et al., 2007); and aspects of depth and figure-ground (e.g., Zhaoping, 2005). It is also tempting to consider whether high level vision structure may similarly be described by a divisive gain control process, as has been hinted in some modeling work on experimental face data (e.g., Bartlett, 2007; Dayan et al., 2003).

The ideas we have described for the case of proximity in space should also apply directly to proximity in time.

There are many commonalities between the spatial and temporal statistical properties of natural scenes that fit the statistical characteristics of a GSM, and it has been seductive to view effects that unfold over time, such as aftereffects, as the temporal equivalent of effects that unfold over space, such as illusions (Clifford et al., 2000; Felsen et al., 2005; Schwartz et al., 2007; Webster, Malkoc, Bilson, & Webster, 2002). However, amidst this general similarity, there are important differences to explore, due to geometry (e.g., Kapadia, Westheimer, & Gilbert, 2000) causality, and timescales (e.g., Kanai & Verstraten, 2005). Finally, a more thorough treatment of perceptual organization should address extensions to multi-modal perceptual organization, as has been suggested in other types of generative statistical models (Kording & Tenenbaum, 2006), and also to additional contextual phenomena beyond illusions and aftereffects, including discriminability and popout (Zhaoping, 2007).

Appendix A

For the GSM model (see also Schwartz et al., 2006b), we assume prior distributions for the two components, the Gaussian and the mixer. The mixer variable is assumed to follow a Rayleigh distribution:

$$p[v] = v\,\exp(-v^2/2). \tag{A1}$$

The Gaussian component is given by:

$$p[g] = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-g^2/2\sigma^2\right), \tag{A2}$$

with mean $\mu = 0$ and standard deviation σ (where, for simplicity, we have set $\sigma = 1$ throughout the main text and in the simulations).

In the paper, we have specifically referred to center and surround filters, with local Gaussian components g_c and g_s; and activations l_c and l_s. Here, for convenience, we develop the more general formulation. We assume there are n filters in the gain pool, with activations $l_1 \ldots l_n$. We label the local Gaussian components $g_1 \ldots g_n$. The single mixer variable is given by v. We would like to estimate the distribution $P[g_1 | l_1 \ldots l_n]$; and the mean $E[g_1 | l_1 \ldots l_n]$.

We start by considering a gain pool with just a single filter ($n = 1$). According to the GSM model:

$$l_1 = v g_1. \tag{A3}$$

We estimate the distribution $P[g_1 | l_1]$. According to Bayes:

$$P[g_1|l_1] \propto P[g_1]P[l_1|g_1]. \tag{A4}$$

The first term, $P[g_1]$, is just a Gaussian distribution, given by Equation A2. The second term is given by $P\left[v = \frac{l_1}{g_1}\right]$, since g_1 is now fixed in Equation A3. Including the constant of proportionality, such that the distribution sums to 1, we obtain:

$$P[l_1|g_1] = P\left[v = \frac{l_1}{g_1}\right] = \frac{|l_1|\exp\left(-\frac{l_1^2}{2g_1^2}\right)}{g_1^2}. \tag{A5}$$

Incorporating the two terms into Equation A4 and finding the constant of proportionality results in:

$$P[g_1|l_1] = \frac{\sqrt{\sigma|l_1|}}{\mathcal{B}\left(-\frac{1}{2},\frac{|l_1|}{\sigma}\right)}\frac{1}{g_1^2}\exp\left(-\frac{g_1^2}{2\sigma^2}-\frac{l_1^2}{2g_1^2}\right). \tag{A6}$$

The constant of proportionality $\mathcal{B}\left(-\frac{1}{2},\frac{|l_1|}{\sigma}\right)$, is the modified Bessel function of the second kind (see also Grenander and Srivastava, 2002).

This can be generalized for more filters ($n > 1$) with a common mixer v, by using Bayes to estimate the local Gaussian component of the first filter:

$$p[g_1|l_1 \ldots l_n] \propto p[g_1]p[l_1 \ldots l_n|g_1]. \tag{A7}$$

We can simplify by taking account that the Gaussian variables $g_1 \ldots g_n$ are mutually independent. Therefore, the values of the other filter responses $l_2 \ldots l_n$ only provide information about the underlying hidden variable v. This results in:

$$p[g_1|l_1 \ldots l_n] \propto p[g_1]p[l_1|g_1]P\left[l_2 \ldots l_n|v = \frac{l_1}{g_1}\right]. \tag{A8}$$

The first and second term we have already estimated above and are identical to the $n = 1$ case. Estimating the third term we obtain:

$$P\left(l_2 \ldots l_n|v = \frac{l_1}{g_1}\right) \propto \left(\frac{g_1}{\sigma|l_1|}\right)^{(n-1)}\exp\left(-\frac{g_1^2}{2\sigma^2}\frac{\sum_{i>1} l_i^2}{l_1^2}\right). \tag{A9}$$

Putting all three terms together and finding the constant of proportionality results in:

$$p[g_1|l_1 \ldots l_n] = \frac{(\sigma|l_1|)^{\frac{1}{2}(2-n)}\left(\frac{|l_1|}{l}\right)^{\frac{1}{2}(2-n)}}{\mathcal{B}\left(\frac{n}{2}-1,\frac{l}{\sigma}\right)}g_1^{(n-3)} \cdot \exp\left(-\frac{g_1^2}{2\sigma^2}\frac{l^2}{l_1^2}-\frac{l_1^2}{2g_1^2}\right), \tag{A10}$$

where $l = \sqrt{\sum_i l_i^2}$. The constant of proportionality $\mathcal{B}$ (n, l), depends on the number of filters n and on l.

We compute the expected value of this distribution, which in the paper simulations amounts to the nonlinear neural response:

$$E[g_1|l_1 \ldots l_n] = \text{sign}\{l_1\}\sigma\sqrt{\frac{|l_1|}{\sigma}}\sqrt{\frac{|l_1|}{l}}\frac{\mathcal{B}\left(\frac{n}{2}-\frac{1}{2},\frac{l}{\sigma}\right)}{\mathcal{B}\left(\frac{n}{2}-1,\frac{l}{\sigma}\right)}. \quad \text{(A11)}$$

Finally, in the paper we label the filter of interest l_c, corresponding to the center; and the other $n - 1$ surround filters (all with equal activations) l_s. Using this formalism, Equation A11 becomes:

$$E[g_c|l_c, l_s] = \text{sign}\{l_c\}\sigma\sqrt{\frac{|l_c|}{\sigma}}\sqrt{\frac{|l_c|}{l}}\frac{\mathcal{B}\left(\frac{n}{2}-\frac{1}{2},l\right)}{\mathcal{B}\left(\frac{n}{2}-1,l\right)} \quad \text{(A12)}$$

where $l = \sqrt{l_c^2 + (n-1)l_s^2 + k}$. Note that we have added k, a small constant, which sets a minimal gain when the filter activations are zero. The constant k acts as a placeholder for model error. In practice, as in divisive gain control frameworks, this constant is a free parameter in our model. In addition, as noted above, we set $\sigma = 1$ in all simulations.

Acknowledgments

This work was funded by the HHMI (OS, TJS); and the Gatsby Charitable Foundation, the BBSRC, the EPSRC and the Wellcome Trust (PD). We are very grateful to Adam Kohn for comments on the manuscript and discussions; and to Colin Clifford, Ruben Coen-Cagli, Isabelle Mareschal, Michael Morgan, Adam Sanborn, Joshua Solomon and Li Zhaoping for discussions. We are particularly grateful to an anonymous reviewer for pointing out that we were plotting aftereffect and not illusion data in an earlier version of the manuscript! We also thank Jitendra Malik and colleagues for making available online the Berkeley Segmentation Dataset.

Commercial relationships: none.
Corresponding author: Odelia Schwartz.
Email: oschwart@aecom.yu.edu.
Address: Albert Einstein College of Medicine, 1410 Pelham Pkwy S. Kennedy Center, Room 921, Bronx, NY 10461, USA.

References

Adelson, E. H. (2000). *Lightness perception and lightness illusions* (pp. 339–351). Cambridge, MA: MIT Press.

Albrecht, D. G., Farrar, S. B., & Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex. *The Journal of Physiology, 347,* 713–739. [PubMed] [Article]

Andrews, D., & Mallows, C. (1974). Scale mixtures of normal distributions. *Journal of the Royal Statistical Society, 36,* 99–102.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61,* 183–193. [PubMed]

Barlow, H. B. (1961). *Possible principles underlying the transformations of sensory messages* (pp. 217–234). Cambridge, MA: MIT Press.

Bartlett, M. (2007). Information maximization in face processing. *Neurocomputing, 70,* 2204–2217.

Bednar, J. A., & Miikkulainen, R. (2000). Tilt aftereffects in a self-organizing model of the primary visual cortex. *Neural Computation, 12,* 1721–1740. [PubMed]

Behrmann, M., Kimchi, R., & Olson, C. (Eds.) (2003). *Perceptual organization in vision: Behavioral and neural perspectives*. Mahwah, NJ: Lawrence Erlbaum Association.

Blake, A., & Zisserman, A. (1987). *Visual reconstruction*. Cambridge, MA: MIT Press.

Blakemore, C., Carpenter, R. H., & Georgeson, M. A. (1970). Lateral inhibition between orientation detectors in the human visual system. *Nature, 228,* 37–39. [PubMed]

Brunswik, E., & Kamiya, J. (1953). Ecological cue-validity of proximity and of other gestalt factors. *Journal of Psychology, 66,* 20–32. [PubMed]

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C. E., Riesenhuber, M., & Poggio, T. (2007). A model of V4 shape selectivity and invariance. *Journal of Neurophysiology, 98,* 1733–1750. [PubMed] [Article]

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17,* 8621–8644. [PubMed] [Article]

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology, 88,* 2547–2556. [PubMed] [Article]

Clifford, C. W. G., & Rhodes, G. (Eds.) (2005). *Fitting the mind to the world: Adaptation and after-effects in high-level vision*. Oxford: Oxford University Press.

Clifford, C. W., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proceedings of the Royal Society of London*

B: Biological Sciences, 267, 1705–1710. [PubMed] [Article]

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. Cambridge, MA: MIT Press.

Dayan, P., Sahani, M., & Deback, G. (2003). Adaptation and unsupervised learning. In S. Becker, S. T. & K. Obermayer (Eds.), *Advances in neural information processing systems 15* (pp. 221–228). Cambridge, MA: MIT Press.

Dragoi, V., & Sur, M. (2000). Dynamic properties of recurrent inhibition in primary visual cortex: Contrast and orientation dependence of contextual effects. *Journal of Neurophysiology, 83,* 1019–1030. [PubMed] [Article]

Durant, S., & Clifford, C. W. (2006). Dynamics of the influence of segmentation cues on orientation perception. *Vision Research, 46,* 2934–2940. [PubMed]

Elder, J. H., & Goldberg, R. M. (2002). Ecological statistics of gestalt laws for the perceptual organization of contours. *Journal of Vision, 2*(4):5, 324–353, http://journalofvision.org/2/4/5/, doi:10.1167/2.4.5. [PubMed] [Article]

Felsen, G., Touryan, J., & Dan, Y. (2005). Contextual modulation of orientation tuning contributes to efficient processing of natural stimuli. *Network, 16,* 139–149. [PubMed]

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. [PubMed]

Geisler, W. S., & Albrecht, D. G. (1992). Cortical neurons: Isolation of contrast gain control. *Vision Research, 32,* 1409–1410. [PubMed]

Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research, 41,* 711–724. [PubMed]

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6,* 721–741.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal population coding of movement direction. *Nature, 233,* 1416–1419. [PubMed]

Gibson, J. J., & Radner, M. (1937). Adaptation, aftereffect, and contrast in the perception of tilted lines. *Journal of Experimental Psychology, 20,* 553–569.

Gilbert, C. D., & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Research, 30,* 1689–1701. [PubMed]

Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., et al. (1999). An anchoring theory of lightness perception. *Psychological Review, 4,* 795–834. [PubMed]

Goddard, E., Clifford, C. W., & Solomon, S. G. (2008). Centre-surround effects on perceived orientation in complex images. *Vision Research, 48,* 1374–1382. [PubMed]

Graham, N., & Sutter, A. (2000). Normalization: Contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision. *Vision Research, 40,* 2737–2761. [PubMed]

Grenander, U., & Srivastava, A. (2002). Probability models for clutter in natural images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23,* 423–429.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9,* 181–197. [PubMed]

Hinton, G. E., & Ghahramani, Z. (1997). Generative models for discovering sparse distributed representations. *Philosophical Transactions of the Royal Society B: Biological Sciences, 352,* 1177–1190. [PubMed] [Article]

Jin, D. Z., Dragoi, V., Sur, M., & Seung, H. S. (2005). Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. *Journal of Neurophysiology, 94,* 4038–4050. [PubMed] [Article]

Kanai, R., & Verstraten, F. A. (2005). Perceptual manifestations of fast neural plasticity: Motion priming, rapid motion aftereffect and perceptual sensitization. *Vision Research, 45,* 3109–3116. [PubMed]

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology, 84,* 2048–2062. [PubMed] [Article]

Karklin, Y., & Lewicki, M. S. (2005). A hierarchical Bayesian model for learning nonlinear statistical regularities in nonstationary natural signals. *Neural Computation, 17,* 397–423. [PubMed]

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97,* 3155–3164. [PubMed] [Article]

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience, 7,* 764–772. [PubMed]

Kording, K., & Tenenbaum, J. B. (2006). Causal inference in sensorimotor integration. In B. Schölkopf, J. Platt, & T. Hoffman (Eds.), *Advances in neural information processing systems* (vol 19, pp. 737–744). Cambridge, MA: MIT Press.

Krüger, N. (1998). Collinearity and parallelism are statistically significant second order relations of complex cell responses. *Neural Processing Letters, 8,* 117–129.

Lehky, S. R., & Sejnowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience, 10,* 2281–2299. [PubMed] [Article]

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature, 387,* 73–76. [PubMed]

Li, W., Thier, P., & Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys. *Journal of Neurophysiology, 83,* 941–954. [PubMed] [Article]

Li, Z. (1999). Visual segmentation by contextual influences via intra-cortical interactions in primary visual cortex. *Network, 10,* 187–212. [PubMed]

Martin, D., Fowlkes, C., Tal, D., & Malik, J. (2001). A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *In Proceedings of the IEEE 8th International Conference on Computer Vision, 2,* 416–423.

Muir, D., & Over, R. (1970). Tilt aftereffects in central and peripheral vision. *Journal of Experimental Psychology, 85,* 165–170. [PubMed]

Mumford, D., & Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics, 42,* 577–685.

Neisser, U. (1967). *Cognitive psychology*. Englewood Cliffs, NJ: Prentice-Hall.

Nundy, S., & Purves, D. (2002). A probabilistic explanation of brightness scaling. *Proceedings of the National Academy of Sciences of the United States of America, 99,* 14482–14487. [PubMed] [Article]

Over, R., Broerse, J., & Crassini, B. (1972). Orientation illusion and masking in central and peripheral vision. *Journal of Experimental Psychology, 96,* 25–31. [PubMed]

Parra, L., Spence, C., & Sajda, P. (2000). Higher-order statistical properties arising from the non-stationarity of natural signals. In *Advances in neural information processing systems* (vol. 13, pp. 786–792). Cambridge, MA. MIT Press.

Portilla, J., Strela, V., Wainwright, M. J., & Simoncelli, E. P. (2003). Image denoising using a scale mixture of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing, 12,* 1338–1351. [PubMed]

Pouget, A., Dayan, P., & Zemel, R. (2000). Information processing with population codes. *Nature Reviews Neuroscience, 1,* 125–132. [PubMed]

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience, 9,* 1421–1431. [PubMed]

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews, Neuroscience, 8,* 522–535. [PubMed]

Schwartz, O., Sejnowski, T., & Dayan, P. (2006a). A Bayesian framework for tilt perception and confidence. In Y. Weiss, B. Schölkopf, & J. Platt (Eds.), *Advances in neural information processing systems* (vol. 18, pp. 1201–1208). Cambridge, MA: MIT Press.

Schwartz, O., Sejnowski, T. J., & Dayan, P. (2006b). Soft mixer assignment in a hierarchical generative model of natural scene statistics. *Neural Computation, 18,* 2680–2718. [PubMed]

Schwartz, O., & Simoncelli, E. (2001a). Natural sound statistics and divisive normalization in the auditory system. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), *Advances in neural information processing systems* (vol. 13, pp. 166–172). Cambridge, MA: MIT Press.

Schwartz, O., & Simoncelli, E. P. (2001b). Natural signal statistics and sensory gain control. *Nature Neuroscience, 4,* 819–825. [PubMed]

Sengpiel, F., Sen, A., & Blakemore, C. (1997). Characteristics of surround inhibition in cat area 17. *Experimental Brain Research, 116,* 216–228. [PubMed]

Series, P., Lorenceau, J., & Frégnac, Y. (2003). The “silent” surround of V1 receptive fields: Theory and experiments. *The Journal of Physiology, 97,* 453–474. [PubMed]

Seung, H. S., & Sompolinsky, H. (1993). Simple models for reading neuronal population codes. *Proceedings of the National Academy of Sciences of the United States of America, 90,* 10749–10753. [PubMed] [Article]

Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22,* 888–905.

Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and gestalt rules. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 1935–1940. [PubMed] [Article]

Simoncelli, E. P. (1997). Statistical models for images: Compression, restoration and synthesis. In *Proceedings of the 31st Asilomar Conference on Signals, Systems and Computers* (pp. 673–678). Pacific Grove, CA: IEEE Computer Society. Available from http://www.cns.nyu.edu/~eero/publications.html.

Simoncelli, E. P., Freeman, W. T., Adelson, E. H., & Heeger, D. J. (1992). Shiftable multi-scale transforms.

IEEE Transactions on Information Theory, 38, 587–607.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38,* 743–761. [PubMed]

Sirosh, J., & Miikkulainen, R. (1997). Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Computation, 9,* 577–594. [PubMed]

Snippe, H. P. (1996). Parameter extraction from population codes: A critical assessment. *Neural Computation, 8,* 511–529. [PubMed]

Solomon, J. A., & Morgan, M. J. (2006). Stochastic recalibration: Contextual effects on perceived tilt. *Proceedings of the Royal Society of London B: Biological Sciences, 273,* 2681–2686. [PubMed] [Article]

Stocker, A., & Simoncelli, E. P. (2006). Sensory adaptation within a Bayesian framework for perception. *Advances in Neural Information Processing Systems, 18,* 1291–1298.

Sundberg, K. A., Mitchell, J. F., & Reynolds, J. H. (2006). Attentional modulation of center-surround interactions in macaque area V4 [Abstract]. *Journal of Vision, 6*(6):746, 746a, http://journalofvision.org/6/6/746/, doi:10.1167/6.6.746.

Teich, A. F., & Qian, N. (2003). Learning and adaptation in a recurrent model of V1 orientation selectivity. *Journal of Neurophysiology, 89,* 2086–2100. [PubMed] [Article]

Wainwright, M. J. (1999). Visual adaptation as optimal information transmission. *Vision Research, 39,* 3960–3974. [PubMed]

Wainwright, M. J., Schwartz, O., & Simoncelli, E. P. (2002). Natural image statistics and divisive normalization: Modeling nonlinearity and adaptation in cortical neurons. In R. Rao, B. Olshausen, & M. Lewicki (Eds.), *Probabilistic models of the brain: Perception and neural function* (chap. 10, pp. 203–222). Cambridge, MA: MIT Press.

Wainwright, M. J., & Simoncelli, E. P. (2000). Scale mixtures of Gaussians and the statistics of natural images. In S. A. Solla, T. K. Leen, & K.-R. Müller (Eds.), *Advances in neural information processing systems* (vol. 12, pp. 855–861). Cambridge, MA: MIT Press.

Wainwright, M. J., Simoncelli, E. P., & Willsky, A. S. (2001). Random cascades on wavelet trees and their use in modeling and analyzing natural imagery. *Applied and Computational Harmonic Analysis, 11,* 89–123.

Webster, M. A., Malkoc, G., Bilson, A. C., & Webster, S. M. (2002). Color contrast and contextual influences on color appearance. *Journal of Vision, 2*(6):7, 505–519, http://journalofvision.org/2/6/7/, doi:10.1167/2.6.7. [PubMed] [Article]

Wenderoth, P., & Johnstone, S. (1988). The different mechanisms of the direct and indirect tilt illusions. *Vision Research, 28,* 301–312. [PubMed]

Wenderoth, P., & van der Zwan, R. (1989). The effects of exposure duration and surrounding frames on direct and indirect tilt aftereffects and illusions. *Perception & Psychophysics, 46,* 338–344. [PubMed]

Wertheimer, M. (1923). Untersuchungen zur lehre von der gestalt, ii [investigations of the principles of gestalt, ii]. *Psychologische Forschung, 4,* 301–350.

Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea. *Vision Research, 30,* 1913–1921. [PubMed]

Williams, C. K. I., & Adams, N. J. (1999). Dynamic trees. In M. S. Kearns, S. A. Solla, & D. A. Cohn (Eds.), *Advances in neural information processing systems* (vol. 11, pp. 634–640). Cambridge, MA: MIT Press.

Wilson, H. R., & Humanski, R. (1993). Spatial frequency adaptation and contrast gain control. *Vision Research, 33,* 1133–1149. [PubMed]

Wolfe, J. M. (1984). Short test flashes produce large tilt aftereffects. *Vision Research, 24,* 1959–1964. [PubMed]

Zetzsche, C., & Röhrbein, F. (2001). Nonlinear and extra-classical receptive field properties and the statistics of natural scenes. *Network, 12,* 331–350. [PubMed]

Zetzsche, C., Wegmann, B., & Barth, E. (1993). Nonlinear aspects of primary vision: Entropy reduction beyond decorrelation. In *International Symposium, Society for Information Display* (vol. XXIV, pp. 933–936).

Zhaoping, L. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron, 47,* 143–153. [PubMed] [Article]

Zhaoping, L. (2007). Colour categories, colour constancy, and lightness perception from information theory. *Perception, 36, ECVP Abstract Supplement.* http://www.perceptionweb.com/abstract.cgi?id=v070251.

Zhaoping, L., & Jingling, L. (2008). Filling-in and suppression of visual perception from context: A Bayesian account of perceptual biases by contextual influences. *PLoS Computational Biology, 4,* e14. [PubMed] [Article]

Zhu, S., & Mumford, D. (1997). Prior learning and Gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19.*